A Population-based Feature Selection and Fuzzy Rule Leukemia form image and Gene Expression Data.

R. Suriyagrace[1], M. Devapriya[2]

Abstract

Genomics is the dominant technique for novel biomarker discovery. However, there are many challenges to handling the image data in the microscopic images extracting significant and robust image features to be mined jointly with genomic (and clinical, etc.). Here chapter gene expression data is combined with image features such as minor axis length, area of the nucleus, the shape of the nucleus, convex area, eccentricity, perimeter, solidarity, orientation, completeness, and the ratio of cytoplasm to nucleus area to detect leukemia effectively.

3085

Introduction

Cancer is the most potent cause of death in India and other countries. Cancer can influence any part of the human body and people of all ages, but the risk for most types of cancer grows with age. Cancers classify the type of cells that constitutes the tumor and, therefore, the tissue presumed to be the origin of the tumor. Such as

Carcinoma: cancer that influences the epithelial tissues that line internal organs. Most known cancers like breast, prostate, lung, and colon cancer come under this category.

Sarcoma: Cancer begins in connective or supportive tissue (e.g., bone, cartilage, fat, muscle, blood vessels).

Leukemia: cancer-related to blood-forming tissue.

Lymphoma: cancers that affect the lymphatic tissue.

Myeloma: cancer that begins in the bone marrow.

Blastoma: cancer that begins in embryonic tissue.

Central nervous system cancers: cancers that begin in the tissues of the brain and spinal cord.

Leukemia is a malignancy (cancer) of blood cells; in leukemia, abnormal blood cells produce in the bone marrow. Generally, leukemia involves the production of abnormal white blood cells -- the cells responsible for fighting infection. However, the abnormal cells in leukemia do not function the same way as normal white blood cells. Instead, the leukemia cells continue to grow and divide, eventually crowding the normal blood cells. As a result, it becomes difficult for the body to fight infections, control bleeding, and transport oxygen. There are different types of leukemia based on how quickly the disease develops and the type of abnormal cells produced. Leukemia is called acute leukemia if it develops rapidly.

Here, in the first step, a microscopic image representation is learned from a collection of images representative of the pathology under investigation. In the second step, an efficient optimisation algorithm called Lion Optimization Algorithm (LOA) introduces to select the most discriminative features from the dataset. The computational complexity because space complexity of the classifier is high by including more number of the classifier. The unique lifestyle of lions and their cooperation characteristics has been the essential motivation for developing this optimisation algorithm. Finally, the selected features train by a Fuzzy deep learning decision system to detect leukemia.

Corresponding author: R. Suriyagrace
Address: [1]Research Scholar, Department of Computer Science, Government Arts College, [2]Assistant Professor, Department of

Computer Science, Government Arts College.

Related work

There are vector machines, k-nearest neighbour, probabilistic neural network, and adaptive neuro-fuzzy inference system. Although they reported good results, a similar threshold setting might fail to provide a consistent performance for microscopic images from different databases.

Choudhary et al. (2017) proposed an automated approach for leukemia detection in human blood samples. In this fast-growing technology, manual counting of WBCs under a microscope is time-consuming, and accuracy depends on the person's skill. This paper proposed an image processing technique for detecting leukemia in a human blood sample. The proposed work overcome the k-means clustering and thresholding method by using image enhancement techniques and some arithmetic operation to segment the nucleus from the white blood cells. Segmentation based on LABcolorr space (luminosity, chromaticity layer a, and chromaticity layer b) colour space will be used to eliminate white blood cells (WBC) from the background. The segmented image is used to calculate the shape-based feature of the nucleus of the WBCs. K- NN classifier has been utilised to classify blast cells from normal lymphocyte cells. The system is applied to 108 images available in the public image dataset for the study of leukemia.

Deshmukh et al. (2018) used automatic segmentation and classification techniques to detect leukemia accurately and within less period. Segmentation scheme segment WBCs into nucleus and cytoplasm, classification is used to classify WBC into various as per different characteristics also, features of nucleus and cytoplasm extracted.

Rajesh &Sathiamoorthy (2018) introduced a new Genetic based KNN preprocessing approach for removing the noise in Leukemia images without affecting the accuracy of an image. This paper integrates the Genetic algorithm and KNN for noise removal and preprocessing Leukemia image datasets. Although these proposed methods are fast in performance, they cannot perform well concerning small-variant cluster segmentation.

Tuba et al. (2019) proposed a method that uses shape and texture features as input vectors for support vector machines optimised by a bare-bones fireworks algorithm. Based on the results obtained on the standard benchmark set, ALL-IDB, the proposed method shows a competitive classification accuracy compared to another state-of-the-art method.

Hegde et al. (2019) presented a comparative study of feature extraction using two approaches for classifying white blood cells. In the first approach, features were extracted using the traditional image processing method, and the second approach employed AlexNet, a pre-trained convolutional neural network, as a feature generator. He refused the Neural network for the classification of WBCs. The results demonstrate that the classification result is slightly better for the features extracted using the convolutional neural network approach than the traditional image processing approach. Furthermore, the average accuracy and sensitivity of 99% were obtained for classifying white blood cells. Hence, any of these methods can be used to classify WBCs depending on the availability of data and rquired resources.

Hegde et al. (2019) automated method for detection of leukemia using an image processing approach. In the present study, 1159 images of different brightness levels and colour shades were acquired from Leishman-stained peripheral blood smears. The SVM classifier was used for the classification of white blood cells into normal and abnormal and also for detecting leukemic WBCs from the abnormal class. The normal white blood cells were classified into five sub-types using an NN classifier. Overall classification accuracy of 98.8% was obtained using the combination of NN and SVM.

Sudha&Geetha(2020)proposed a new leukocyte segmentation framework that first locates and then segments leukocytes from peripheral blood images. Here, the locations of the leukocytes are first identified using a novel edge strength cue (ESC), and later, the grab cut model is deployed to obtain the segmentation of the leukocytes. The novelty lies in how the location of the leukocytes is detected, which improves the leukocyte segmentation accuracy. The experimental evaluation is performed on ALL-IDB1, Television, and LISC datasets for leukocyte segmentation based on detecting the ESC location. Experimental results are evaluated using precision, recall, and F-score measures. The proposed method outperforms the state-of-the-art techniques. Additionally, the computation time of the proposed method is analysed and presented in the study.

Proposed methodology

Radio-genomics pipeline used in the examination of the association between imaging features and gene

NEUROQUANTOLOGY | OCTOBER 2022 | VOLUME 20 | ISSUE 12 | PAGE 3085-3097| DOI: 10.14704/NQ.2022.20.12.NQ77309

R. Suriyagrace et al / A Population-based Feature Selection and Fuzzy Rule Leukemia form image and Gene Expression Data.

signatures in patients with leukemia cancer. First, microscopic samples were acquired then the enriched feature extraction process was done. Then the feature selection using Lion Optimization Algorithm (LOA) is applied. For microscopic leukemia samples, RNA was extracted, and gene expression was measured using micro-array technology. The ferrule-based based decision support system is used for leukemia detection with the selected radiomic features. A similar technique was used for contra-lateral images to evaluate whether the associations were leukemia-specific.
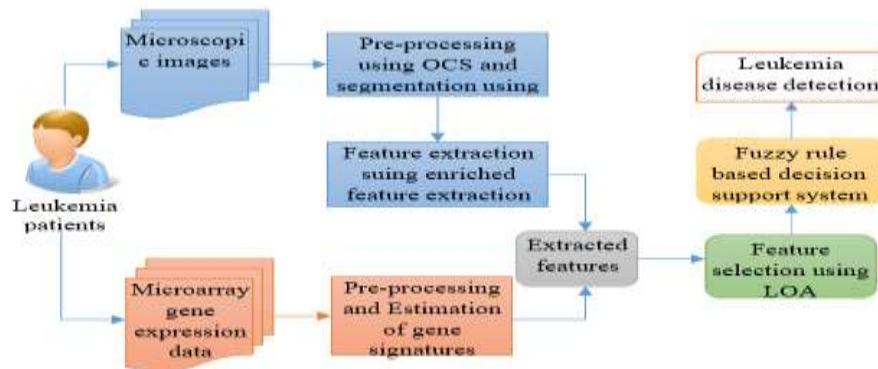

Figure 3.1. Research Workflow of the proposed leukemia detection using the image and gene expression data

Input Gene expression data set

In this section, five datasets from NCBI Gene Expression Omnibus (GEO) (Barrett et al., 2012) cancer gene expression datasets were used for the performance analysis of the algorithm. This study also used the leukemia gene expression data collected by Golub et al. (1999). This data set comprised 72 bone marrow samples, of which 38 were B-cell ALL, 9 were T-cell ALL,d 25 were AML samples.

Data pre processing

The training data set consisted of gene expression profiles for 37 patients. Each profile comprised 7129 gene expression values. The preprocessing was carried out along the same lines as Peng et al. (2014). Endogenous control genes were eliminated from consideration. Genes with "absent" calls across samples were eliminated from consideration. Genes with less than 2.5 fold change across samples were eliminated from consideration. Genes with expression values of less than 20 and greater than 16000 were thresholded to 20 and 16000, respectively. The remaining gene expressions were normalized so that the expression levels were in the range of −1 to 1.

Image Dataset

This dataset contains 12,500 augmented images of blood cells (JPEG) with accompanying cell type labels (CSV). There are approximately 3,000 images for each of 4 different cell types grouped into 4 different folders (according to cell type). The cell types are Eosinophil, Lymphocyte, Monocyte, and Neutrophil.

3087

Image Pre-processing Using OCS

The preprocessing of microscopic blood images by contrast stretching (CS) enhances the global uniformity, local sensitivity and geometry of the blood cells. It is performed by morphological addition and subtraction operation illustrated in algorithm provides less computation complexity. The proposed algorithm improves the contrast of the blood cell, which reflects in the projection of nuclei in the WBCs.

Refined Segmentation Using K-means and HMRF clustering

This section proposed a method to analyse the aspirate smear images, which first performs segmentation of the cells by k-means cluster, then builds cell image representing
model by HMRF (Hidden-Markov Random Field), estimates model parameters through the probability of EM (expectation maximisation), carries out convergence iteration until optimal value, and finally achieves second stage refined segmentation.

Enriched Feature Extraction Process

Feature extraction plays an important role in the performance of an automatic white blood cell classification system. Most of the existing methods adopt the following features such as geometrical features (e.g., area, radius, perimeter, convex area, major axis length, compactness, and orientation), textural features (e.g., momentum, contrast, entropy, and skewness), and colour features (e.g., colour distribution and histogram). In this phase, these four kinds of features are extracted for classification. The first kind of feature is the geometrical feature consisting of the area feature, area, length feature, Lengthvar, and the compactness feature, Comp. The area feature, area, is the amount numbers that belong to the segmented cell region.

Feature selection using Ant-Lion Optimization Algorithm (ALOA)

ALOA algorithm is adopted for image feature selection. The artificial ant-lions are divided into several classes. Ant-lions of each class traverse on a feature graph where (edge) the pride is life represents features, and the edge connecting two adjacent nodes represents one of the discrete values of the feature. At each step, the Lion on one node (feature) selects an edge connecting another node (feature) based on the pheromone and heuristic information assigned to this edge. After each iteration, each Ant-Lion constructs a solution path which represents pairs of features and their discrete value so as to form the classification rules of each class while maintaining fairly high classification accuracy.

Ant-Lion Optimization Algorithm

The Ant-Lion Optimization Algorithm (ALOA), which was presented by Wang et al. (2012) and Rajakumar (2012), is based on the Ant-Lion's social behaviour. ALOA is a meta-heuristic algorithm which is a part of stochastic optimization. Meta-heuristic algorithms are able to generate different solutions for the problem in each run. Ant-Lion has unique social behaviour, so it is the strongest mammal in the world. Pride generation is the initiating process of the lion algorithm, which is similar to the initialization steps of most of the evolution and swarm-based optimisation algorithms. Mating is the most responsible process for deriving new lions (called cubs) from the parent lions after subjecting them to fertility evaluation, which is the second process.

Territorial defence and territorial takeover are identified as unique processes over other optimization algorithms, as they are the explicit inspiration of the Lion's social behaviour. These two steps play the primary roles in guiding the algorithm in determining the optimal solutions from a huge search space. The termination process of the lion algorithm is problem-dependent, and hence that could be either based on a number of iterations/generations or based on the optimal of the obtained solutions. More details about individual processing stages are described in (Rajakumar, 2012), while their concise description is given further. The steps to be followed to simulate the lion algorithm are as follows in Table 3.1.

3088

Table 3.1.The algorithm of LOA

Step 1: Initialize l^{male}, l^{nomad} and l^{nomad}

Step 2: Calculate $f(l^{male})$, $f(l^{nomad})$ and $f(l^{nomad})$ //the f(.) is calculated as $l^{optimal} =$

NEUROQUANTOLOGY | OCTOBER 2022 | VOLUME 20 | ISSUE 12 | PAGE 3085-3097| DOI: 10.14704/NQ.2022.20.12.NQ77309

R. Suriyagrace et al / A Population-based Feature Selection and Fuzzy Rule Leukemia form image and Gene Expression Data.

$$\arg \min_{l_i \in (l_i^{min}, l_i^{max})} f(l_1, l_2, ..., l_n) \quad \forall n \geq 1$$

Step 3: Set $f^{ref} = f(l^{male})$ and $N_g = 0$ // N_g is the number of generations, reference fitness

Step 4: Store l^{male} and $f(l^{male})$ then do fertility evaluation and also perform mating and obtain cubpool

Step 5: Perform gender clustering and obtain $l^{male-cub}$ and $l^{female-cub}$

Step 6: Initialize cub age as zero and execute cub growth function and perform territorial defense; if defense result 0, go to step 4

Step 7: If cub age < age mat, go to step 6

Step 8: Perform territorial takeover and obtain updated l^{male} and l^{female}

Step 9: Increase N_g by one, if the termination criteria are not met, go to step 4, otherwise terminate the process.

LOA searches for optimal solutions based on two lion behaviour proposed by Rajakumar (2012):
Territorial Defense
Territorial Takeover

Proposed Feature selection using LOA

Using LOA to improve feature selection in image processing by searching and discovering hidden relationships between features. LOA needs some modification to apply it to the feature selection problem. To represent individuals (features) in image processing, we chose the locus-based adjacency encoding scheme, which depends on a genetic algorithm (). The lion position represented as l consists of n elements ($l1, l2, ...ln$). A value j in the range [1,···, n] is assigned to each element n . Also, a value j is assigned to the i element. The assignment between i and j simulates the relation between individual i and individual j, so the two individuals are in one range. The position of the Lion represents a solution in the search space, and the fitness value of each Lion is calculated. The cost (fitness value) of each Lion is computed by evaluating the cost function as:

$$fitness \ value \ of \ lion = f(l1, l2, ..lNvar)$$

The steps used from LOA in the feature selection problem:

Initialisation. The population is randomly generated over the solution space, and the position of the lions is saved in the matrix. The fitness value of each Lion is calculated by evaluating the objective function, sorted and saved in the matrix as in Eq.(3.1).

Mating. It is a process of providing an opportunity to produce new best solutions from the existing solutions, including crossover and mutation. Killing sick/weak cubs ensures that the derived solutions are the best.

Territorial Defence. Compare between fitness value of the resident lion $lmale$ and nomadic Lion ($lnomad$), and if the fitness value of the nomadic Lion is better than the fitness value of resident one, the resident Lion replaced with the nomadic Lion. The algorithm of territorial defencebehaviour is in Table 3.2.

Territorial Takeover. Sort all resident lions (males and cubs) in the pride according to their fitness. Drive out the weakest male from the pride and become a nomad. They remind males to become resident males. The algorithm of territorial takeover behaviour is given in Table 3.3.

3089

Fuzzy rule (FL) based Decision Support System (DSS) for Leukemia Detection

FL-based DSS and examples of studies in medicine were presented. Important databases, such as Elsevier, Wiley and Springer, were scanned in the research process. The selected search criteria were determined as "FL", "Diagnosis with FL", "DSS in medicine", and so on. Fuzzy systems are one of the most active research fields in recent years because it has many benefits in solving complicated non-linear system modelling. Fuzzy systems' most important advantages are the simplicity for the obvious knowledge representation in the form of if-then rules, a mechanism of reasoning in understandable human terms, the capacity to take linguistic information from human experts and combine it with numerical information, and the capability of approximating complex non-linear functions with simple models.

(3.1)

Unlike classic modelling, where a single model is used to describe the global behaviour of a system, fuzzy rule-based modelling is essentially a multi-model approach in which individual rules are combined to describe the global behaviour of the system Soria et al. (2013). The overview of DSS

with FL is located in Figure 3.3. From figure 3.3, basically, the operation of clinical DSS begins with planning, and the next step is preprocessing of data.

Data preprocessing can be done, then the data is processed in DSS, and a system decision (disease diagnosis) is obtained.
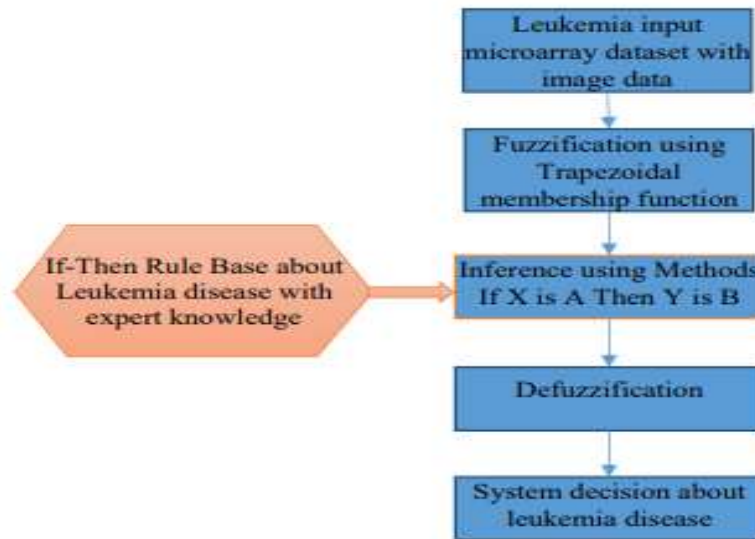

Figure 3.3.Block Diagram of FL-based DSS for leukemia detection

First of all, researchers are required to plan their research and move in this direction expediently, as in many statistical analysis methods. Accurately and reliably obtained data in accordance with research subjects directly affects the accuracy of the analysis and the quality of research. The data analysed in the FL-based DSS in the medical field may have various parameters associated with the diseases and disease inputs. This condition may change according to the model the researchers will perform. Fuzzy systems are knowledge-based systems that are configured using FL theory. FL-based DSS basically consists of three components. These are fuzzification, fuzzy inference engine and defuzzification.

Fuzzy inference process

In FL, the process that formulates the FL-based mapping from a given input to output is known as the fuzzy inference process. It has five steps, and they are fuzzification of the input variables, application of the fuzzy operator (AND or OR) in the antecedent, the implication from the antecedent to the consequent, aggregation of the consequents across the rules, and defuzzification.

Fuzzify inputs

First, the degree of membership of each of the inputs to each of the proper fuzzy sets to which they belong is determined by means of membership functions. The input of the FL Toolbox is restricted to the universe of discourse of the input variable, and it is always a crisp numerical value, whereas its output is a fuzzy degree of membership in the qualifying linguistic set (always the interval between 0 and 1).

Apply fuzzy operator

The degree to which each part of the antecedent has been satisfied for each rule can be known once the inputs have been fuzzified. If more than one part is present in the antecedent of a given rule, a number that represents the result of the antecedent for that rule is obtained by applying the fuzzy operator. This number is then applied to the output function. Two or more membership values from fuzzified input variables are given as input to the fuzzy operator, and a single truth value is obtained as the output.

Apply implication method

The rule's weight must be checked before the implication method is applied. The significance associated with each rule (a number between 0 and 1) is applied to the number given by the antecedent. Because this weight is commonly 1, it does not affect the implication process.

Aggregate all outputs

Combining the fuzzy sets that represent the outputs

of each rule into a single fuzzy set is known as aggregation. Just before defuzzification, the fifth and final step, collection occurs only once for each output variable. The list of truncated output functions returned by the implication process for each rule is given as input to the aggregation process. One fuzzy set for each output variable is provided as the output.

Defuzzify

The defuzzification process accepts a fuzzy set (the output of the aggregate step) as input and outputs a single number. Fuzziness is useful for rule evaluation during the intermediate stages, but generally, a single number output is desired in the final grade for each variable. Since a range of output values is present in the aggregate fuzzy set, defuzzification is utilised to resolve a single output value from the set.

Performance Evaluation

The results of applying the proposed method show satisfactory classification of cells and high values of statistical evaluation parameters. For example, the development of variety in four types of leukemia images such as AML, ALL, CLL and CML, is shown in figures 4.1 to 4.7, respectively.

(a)   (b)   (c)
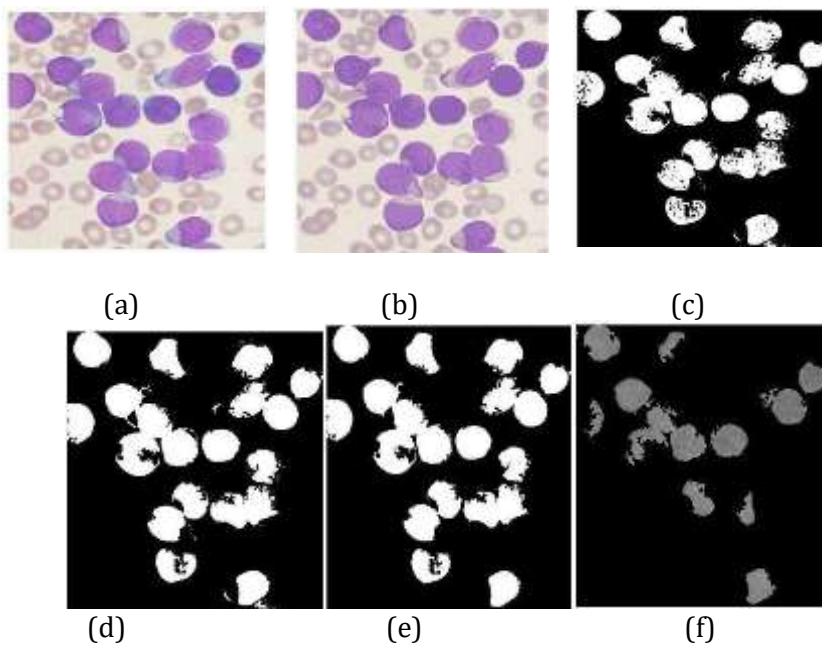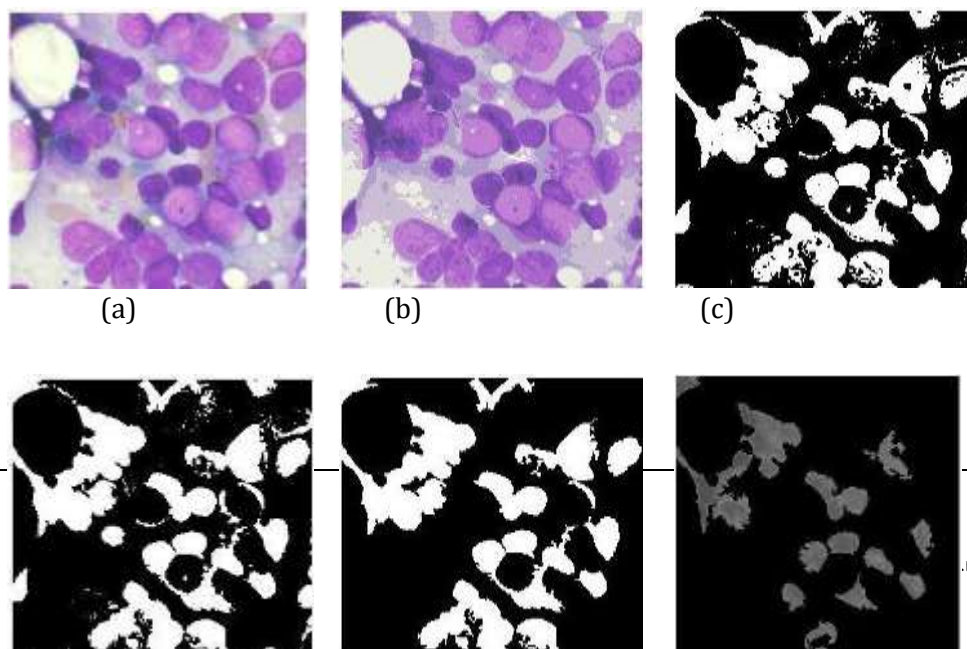
(d)   (e)   (f)

Figure 4.4. Results of the proposed algorithm for ALL (a) Actual image, (b) Enhanced images by OCS, (c) segmented nuclei results of K-means output (middle of the iteration), segmented nuclei results of K-means output (end of the iteration)(e) segmented nuclei results of K-means –HMRF output (middle of the iteration) and (f) segmented nuclei results of K-means –HMRF output (end of the iteration)

(a)   (b)   (c)

(d)   (e)   (f)

Figure 4.5. Results of the proposed algorithm for AML (a) Actual image, (b) Enhanced images by OCS, (c) segmented nuclei results of K-means output (middle of the iteration), segmented nuclei results of K-means output (end of the iteration) (e) segmented nuclei results of K-means –HMRF output (middle of the iteration) and (f) segmented nuclei results of K-means –HMRF output (end of the iteration)
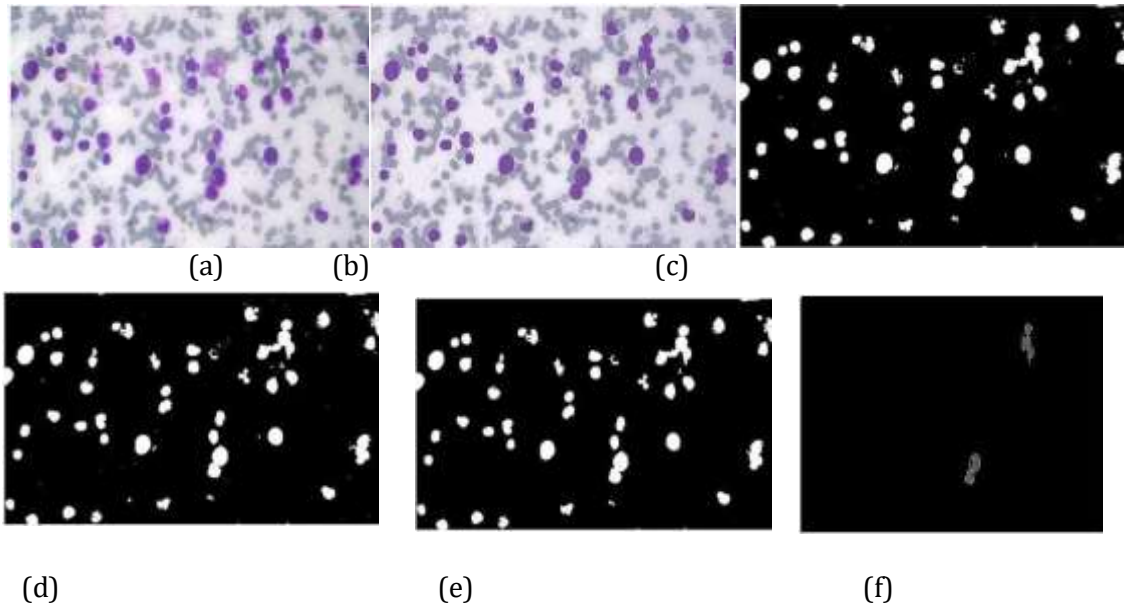

(a)   (b)   (c)

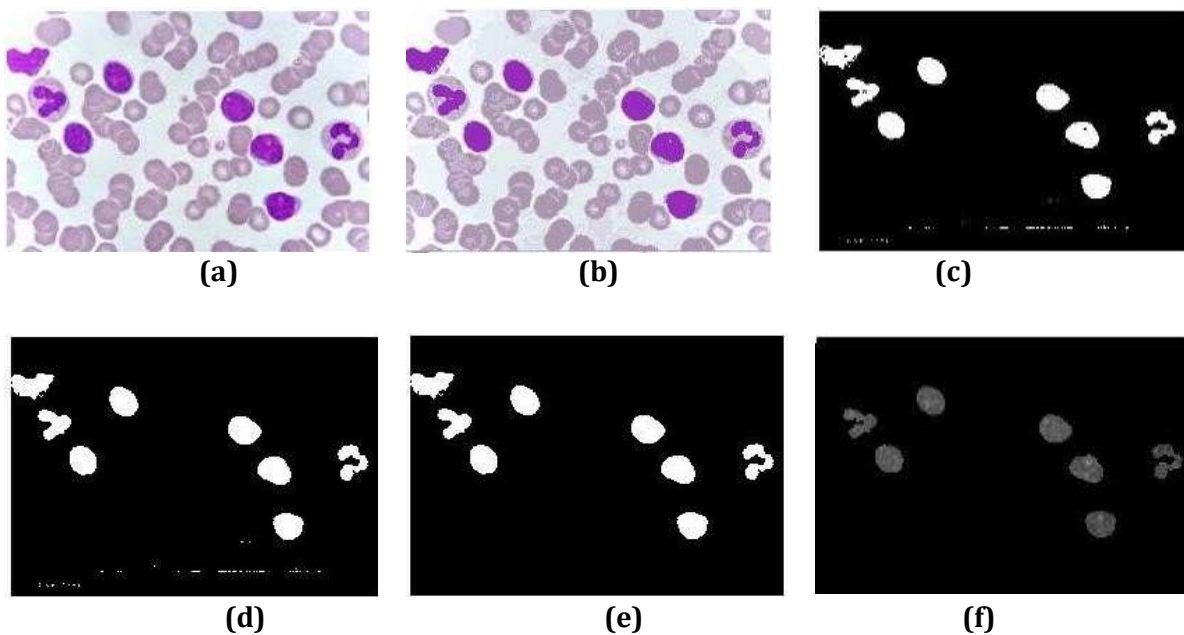(d)   (e)   (f)

3092

Figure 4.6. Results of the proposed algorithm for CL€L (a) Actual image, (b) Enhanced images by OCS, (c) segmented nuclei results of K-means output (middle of the iteration), segmented nuclei results of K-means output (end of the iteration) (e) segmented nuclei results of K-means –HMRF output (middle of the iteration) and (f) segmented nuclei results of K-means –HMRF output (end of the iteration)

(a)   (b)   (c)

(d)   (e)   (f)

Figure 4.7. Results of the proposed algorithm for CML (a) Actual image, (b) Enhanced images by OCS, (c) segmented nuclei results of K-means output (middle of the iteration), (d) segmented nuclei results of K-means output (end of the iteration) (e) segmented nuclei results of K-means –

R. Suriyagrace et al / A Population-based Feature Selection and Fuzzy Rule Leukemia form image and Gene Expression Data.

HMRF output (middle of the iteration) and (f) segmented nuclei results of K-means –HMRF output (end of the iteration)

The performance of the classifiers of exiting marker-based segmentation (MBS), grey level co-occurrence matrix (GLCM) based feature extraction. Probabilistic principal component analysis (PPCA) based feature reduction and random forest (RF) based classifier (MBS+GLCM+PPCA+RF) proposed by Mishra et al. (2017)and presented Combined Image and Gene data with LOA-based OCS+ RS + EFE+ FRDSS (IG+OCS+ RS + EFE+ LOA+ FRDSS) are evaluated by these parameters: precision, recall, f-measure, segmentation accuracy and classification accuracy.
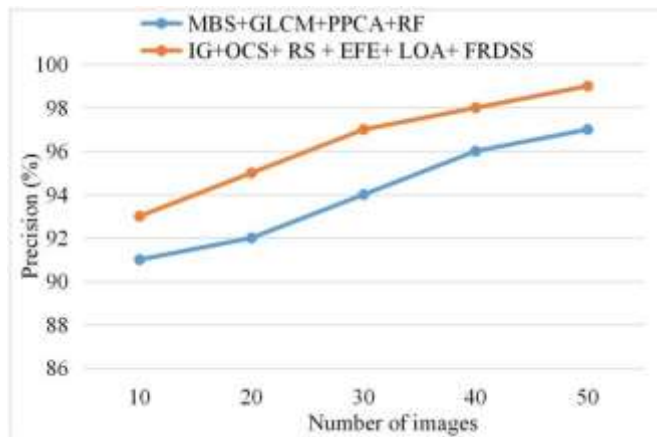
Precision comparison results

Figure 4.8. Precision performance comparison

Figure 4.8 shows the precision comparison results of the microscopic image dataset between existing MBS+GLCM+PPCA+RF and proposed IG+OCS+ RS + EFE+ LOA+ FRDSS based on the number of images. The figure shows that the proposed method can obtain a high precision rate compared to the existing process. It effectively gets the leukemia nucleus precisely with a high precision rate of 99%. When comparing the precision with the current MBS+GLCM+PPCA+RF method providing low precision rates of 97%, which is lower than the IG+OCS+ RS + EFE+ LOA+ FRDSS. The numerical results of Precision rate are shown in Table 4.1.

Table 4.1 The numerical results of Precision rate

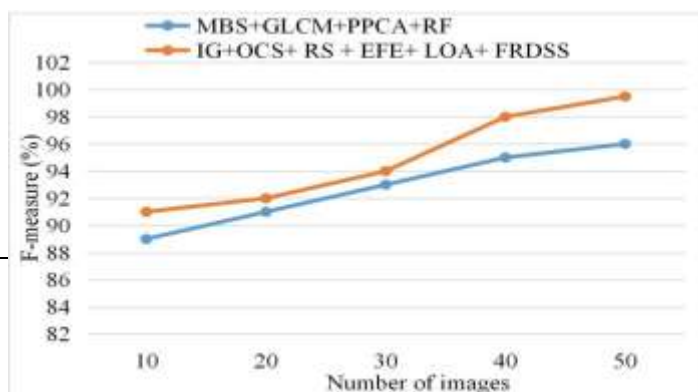| Number of images	MBS+GLCM+PPCA+RF	IG+OCS+ RS + EFE+ LOA+FRDSS
10	91	93
20	92	95
30	94	97
40	96	98
50	97	99

F-measure Result Comparison

Figure 4.9. F-measure performance comparison

Figure 4.9shows the F-measure comparison results between proposed IG+OCS+ RS + EFE+ LOA+ FRDSS and existing MBS+GLCM+PPCA+RF based on the number of images. The proposed method has a high value of F-measure 99.5%, which has the advantage over the existing algorithm` when the image noise is not negligible. On the other hand, when comparing the F- measure rate with the current method, MBS+GLCM+PPCA+RF provides fewer rates of 96%, which indicates the proposed work can give better segmentation results of leukemia nucleus with high quality. The numerical results of F-measure rate are shown in Table 4.2.

Table 4.2. The numerical results of F-measure rate

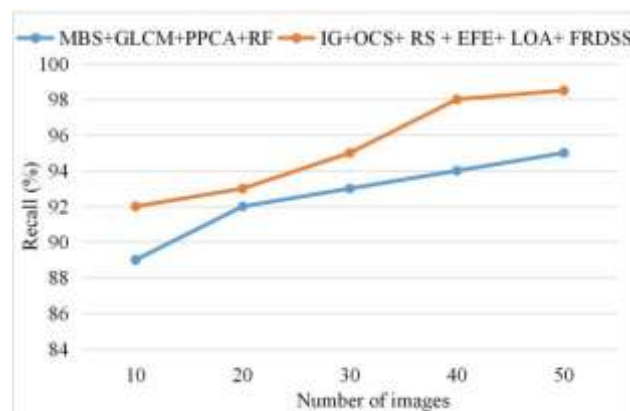| Number of images	MBS+GLCM+PPCA+RF	IG+OCS+ RS + EFE+ LOA+ FRDSS
10	89	91
20	91	92
30	93	94
40	95	98
50	96	99.5

Recall Result Comparison

Figure 4.10. Recall performance comparison

Figure 4.10 shows the recall comparison results between proposed IG+OCS+ RS + EFE+ LOA+ FRDSS and existing MBS+GLCM+PPCA+RF based on the number of images. The proposed MBS+GLCM+PPCA+RF method has a high-value recall rate of 98.5%. When comparing the recall rate among the existing techniques, MBS+GLCM+PPCA+RF provides a recall rate of 95%. The proposed model used the HMRF model for sophisticated segmentation and more feature extraction techniques; thus, accurate and robust segmentation of the nucleus can be achieved. The numerical results of the recall are shown in Table 4.3.

Table 4.3. The numerical results of recall

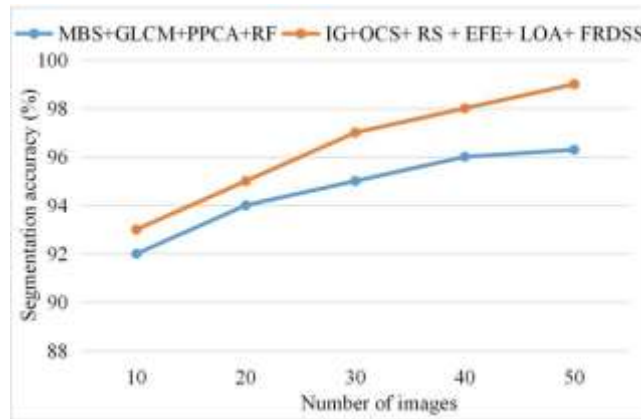| Number of images	MBS+GLCM+PPCA+RF	IG+OCS+ RS + EFE+ LOA+ FRDSS
10	89	92
20	92	93
30	93	95
40	94	98
50	95	98.5

Segmentation Accuracy comparison

Figure 4.11. Result of Accuracy

From the above Figure 4.11, the graph explains the segmentation accuracy comparison for the prediction of nucleus between IG+OCS+ RS + EFE+ LOA+ FRDSS and existing MBS+GLCM+PPCA+RF. When the number of images increases according to the segmentation accuracy value is increased linearly. This graph shows that the proposed IG+OCS+ RS + EFE+ LOA+ FRDSS effectively select the cluster centre with a high segmentation accuracy rate of 99% for nucleus detection. However, the previous methods, such as MBS+GLCM+PPCA+RF, attain an accuracy rate of 96.29%, much lower than the IG+OCS+ RS + EFE+ LOA+ FRDSS. The numerical results of segmentation accuracy are shown in Table 4.4.

Table 4.4. The numerical results of segmentation accuracy

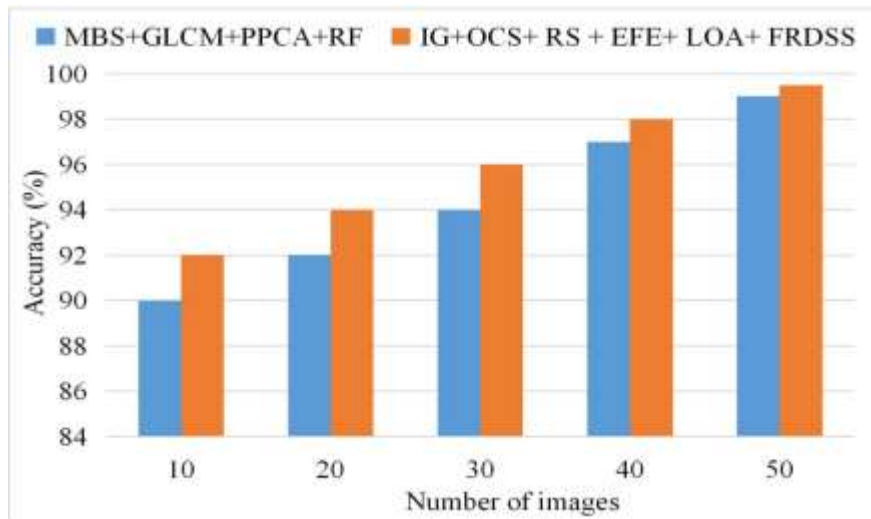| Number of images	MBS+GLCM+PPCA+RF	IG+OCS+ RS + EFE+ LOA+ FRDSS
10	92	93
20	94	95
30	95	97
40	96	98
50	96.29	99

Classification Accuracy comparison

Figure 4.12. Result of Accuracy

From the above Figure 7.12, the graph explains the accuracy comparison for the prediction of leukemia with nucleus segmentation between proposed OCS+RS+EFE+ FRDSS and existing MBS+GLCM+PPCA+RF. When the number of images increases according to the accuracy value is increased linearly. This graph shows that the proposed OCS+RS+EFE+ FRDS Effectively selects the nucleus and detects the leukemia disease with a high accuracy rate of 99.5%. However, the previous method attains an accuracy rate of 99.004%, much lower than the IG+OCS+ RS + EFE+ LOA+ FRDSS. The numerical results of accuracy are shown in Table 4.5.

3096

Table 4.5. The numerical results of accuracy

Number of images	MBS+GLCM+PPCA+RF	IG+OCS+ RS + EFE+ LOA+ FRDSS
10	90	92
20	92	94
30	94	96
40	97	98
50	99.004	99.5

Conclusion

The fuzzy rule-based decision systems utilized for detecting leukemia disease, the quality of health organizations plays an important role in indirectly improving. Then the enriched feature extraction is done; after that, LOA is utilizing to determine significant and optimal features. These association patterns and optimal values provide pertinent knowledge for leukemia diagnosis. Ultimately, FL was used to model the decision module to perform the actual diagnosis. The experimental results have demonstrated the effectiveness of the proposed feature selection and fuzzy-based decision support system in Leukemia disease diagnosis with an accuracy rate of 99.5%.

Reference

Abualigah, Laith, Aldulaimi, Akram, Al Shinwan, Mohammad, Khasawneh, Ahmad &Alabool, Hamzeh, Diabat, Mofleh& Shehab, Mohammad. (2020). Optimisation Algorithms to Solve Feature Selection Problem: A Review. International Journal of Science and Applied Information Technology. 8. 10.30534/ijsait/2019/098620198.

Choudhary, R. R., Sharma, S., & Meena, G. (2017, October). Detection of Leukemia in Human Blood Samples Through Image Processing. In International Conference on Next Generation Computing Technologies (pp. 824-834). Springer, Singapore.

Deshmukh, P., Jadhav, C. R., & Rani, N. U. (2018). Automatic White Blood Cell Segmentation for Detecting Leukemia. In Information and Communication Technology for Sustainable Development (pp. 385-392). Springer, Singapore.

Dutta, M., Karmakar, S., Banerjee, P., &Ghatak, R. (2020). Detection of Leukemia in Blood Samples Applying Image Processing Using a Novel Edge Detection Method. In Proceedings of the Global AI Congress 2019 (pp. 1-16). Springer, Singapore.

Hegde, R. B., Prasad, K., Hebbar, H., Singh, B. M. K., & Sandhya, I.

(2019). Automated Decision Support System for Detection of Leukemia from Peripheral Blood Smear Images. Journal of Digital Imaging, 1-14.

Hegde, R.B., Prasad, K., Hebbar, H. et al. Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: a study. Australas Phys Eng Sci Med 42, 627–638 (2019).

Hegde, R.B., Prasad, K., Hebbar, H. et al. Image Processing Approach for Detection of Leukocytes in Peripheral Blood Smears. J Med Syst 43, 114 (2019).

Meetei, K. T. (2014). A survey: swarm intelligence vs genetic algorithm. International Journal of Science and Research (IJSR), 3, 231-235.

Peng, H. Y., Jiang, C. F., Fang, X., & Liu, J. S. (2014). Variable selection for Fisher linear discriminant analysis using the modified sequential backward selection algorithm for the microarray data. Applied Mathematics and Computation, 238, 132- 140.

Rajakumar BR (2014) Lion algorithm for standard and large scale bilinear system identification: a global optimisation based on Lion's social behaviour. In: 2014 IEEE congress on evolutionary computation (CEC), pp 2116–2123

Rajesh, M. B., &Sathiamoorthy, S. (2018). Classification of Leukemia Image Using Genetic Based K-Nearest Neighbor (G-KNN). Asian Journal of Computer Science and Technology, 7(2), 113-117.

Soria D, Garibaldi JM, Green AR, Powe DG, Nolan CC, Lemetre C, Ball GR, Ellis IO 2013. A quantifier-based fuzzy classification system for breast cancer patients. ArtifIntell Med, 58: 175-184

Sudha, K., & Geetha, P. (2020). Leukocyte segmentation in peripheral blood images using a novel edge strength cue-based location detection method. Medical & Biological Engineering & Computing, 1-14.

T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, et al., NCBI GEO: Archive for functional genomics data sets- update, Nucleic Acids Research 41 (D1) (2012) D991{D995.

Takiar, R., Nadayil, D., & Nandakumar, A. (2010). Projections of number of cancer cases in India (2010-2020) by cancer groups. Asian Pac J Cancer Prev, 11(4), 1045-9.

Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science. 1999;286:531–7.

H. Bauer, H.H. de longh and I. Di Silvestre, Lion (Panthera leo) social behaviour in the West and Central African savannah belt, Mammalian Biology-ZeitschriftfUrSaugetierkunde, Elsevier, Vol. 68(4), pp. 239-243, 2003.

WANG Bo, JIN Xiaoping and CHENG Bo, "Lion pride optimiser: An optimisation algorithm inspired by lion pride behaviour", Science China Information Sciences, Springer, Vol. 55(10), pp. 2369-2389, 2012.

B.R.Rajakumar, "The Lion's Algorithm: A New Nature-Inspired Search Algorithm", Procedia Technology, Elsevier, Vol. 6, pp. 126-135,2012.

WHO. (2015). Media centre. Available: http://www.who.int/mediacentre/factsheets/fs297/en/